

# National Center for Computational Sciences Snapshot

## The Week of September 24, 2007

### **Simulation Helps Unlock the Secrets of Protein Folding**

*Water determines shape, shape determines function*

A team led by biophysicist Jeremy Smith of the University of Tennessee (UT) and Oak Ridge National Laboratory (ORNL) has taken a significant step toward unraveling the mystery of why proteins fold into unique, three-dimensional shapes.

Using ORNL's Cray XT4 Jaguar supercomputer as well as computer systems in Italy and Germany, the team revealed a driving force behind protein folding involving the way its constituents interact with water. The team's results are being published in the September 18 edition of the Proceedings of the National Academy of Sciences.

Proteins are the workhorses of the body, taking on a wide variety of tasks. They fight infections, turn food into energy, copy DNA, and catalyze chemical reactions. Insulin is a protein, as are antibodies and many hormones.

Nevertheless, scientists are still deciphering how proteins work. A protein is a string of amino acids, but what it does is determined by the shape it takes, which is itself determined by the sequence of the amino acids. Like a piece of biological origami, the protein folds itself into the form necessary to carry out its job. Without the shape the protein would be worthless.

"Understanding the mechanism by which proteins fold up into unique three-dimensional architectures is a holy grail in molecular biology," explained Smith, who holds the first Governor's Chair at UT and ORNL. "If you give me the sequence of amino acid building blocks in the protein, I cannot tell you what the structure would be. If I had been able to do that with a computer a while ago, the work behind about a dozen Nobel prizes—those awarded for experimental work on protein structure determination— would not have been necessary."

Working on a smaller chain of amino acids known as a peptide, the group showed that the folding is determined largely by the aversion of some parts of the peptide to water. Areas that shun water are said to be hydrophobic, and the team's results show how the way in which water wets these hydrophobic areas determines the ultimate shape and behavior of the peptide.

In particular, the team determined that small hydrophobic areas of the peptide, up to the size of a water molecule, induce different behavior in water than larger hydrophobic areas, and that this difference is crucial for the folding. This insight builds on the work of another team, based at the University of California–Berkeley.

"David Chandler and his colleagues at Berkeley have a theory stating that hydrophobicity is qualitatively different on different length scales," said Smith. "If you have small

hydrophobic molecules or groups that are roughly the size of a water molecule, the water doesn't seem to be too bothered by these groups. When you get big enough hydrophobic entities that cover several water molecules of length, however, then the water molecules have a problem with that. They can't cloak themselves around the hydrophobic surface anymore, and there is a dewetting or drying effect as they move away from the surface.

"Our simulations have shown that Chandler's theory works for peptides, and, moreover, that the drying effect determines which structure the peptide adopts."

Smith said his team's achievement was made possible by the computing muscle brought to the task by high-performance computing, noting that Jaguar is currently rated the second most powerful computing system in the world. Smith also noted that his team will need increasingly powerful supercomputers for additional simulations. While the team so far has been able to simulate peptides for a matter of microseconds, it must eventually be able to increase that time a thousand-fold, to milliseconds, and simulate proteins that are ten to 100 times as large as the peptides.

"The runs were a couple of microseconds, which was adequate for the peptide that was simulated," Smith explained. "But the team is looking forward to increased computing capacity as it moves forward. The technique used is molecular dynamics simulation and it needs high-performance leadership supercomputing to reach the length and timescales needed to fold a complete functional protein in the computer. With the projected capability improvements in Jaguar over the next couple of years we will soon be approaching that goal."

Smith made it clear that the achievement would represent a watershed in the field.

"When we do eventually find out how to calculate protein structure from sequence" he said, "then a major revolution will come upon us, as we will have the basis to move forward with understanding much of biology and medicine, and the applications will range from rationally designing drugs to fit clefts in protein structures to engineering protein shapes for useful functions in nanotechnology and bioenergy."

### **NCCS Keeps America Competitive**

*NCCS research exemplifies "America COMPETES"*

The National Center for Computational Sciences (NCCS) recently had an opportunity to show members of Tennessee's congressional delegation how the center is using high-performance computing to boost American scientific and industrial competitiveness.

On August 30, Senator Lamar Alexander and U.S. Representative Zach Wamp visited the NCCS to see firsthand how it is advancing the recent landmark legislation entitled "America Competes." The legislation promotes scientific innovation through increased federal funding for research and development and various educational initiatives.

Sean Ahern, visualization lead for the NCCS Scientific Computing Group, outlined the organization's role to Alexander, Wamp, and several members of the press (including a representative from the New York Times). He pointed to projects that exemplify the goals of the legislation, including two that use the NCCS's supercomputers to study the breakdown of organic materials, such as grasses and cellulose, for the possible production of alternative fuels and better drug design.

Ahern also discussed the recent work of ORNL's Ed Uberbacher, whose simulations are shedding light on the fundamental structure of Alzheimer's disease and revealing how certain drugs interact with the disease. These simulations are more advanced than anything currently being done by drug companies because of the enormous computing power available at the NCCS, said Ahern.

Finally, Ahern discussed cooperation between the NCCS and researchers at General Motors, who have recently used the high-performance computing resources at the NCCS to study the possibility of recycling exhaust and converting it to electricity.

The "America Competes" legislation is a response to a recent study, said Ahern, which concluded that America today is scientifically and industrially competitive, primarily due to investments made 30 to 40 years ago. The increased investment provided by the legislation aims to further America's leadership in the international scientific and industrial communities.

### **NCCS Boosting Power to Building for Petascale System**

*Power doubled to accommodate new supercomputer*

ORNL's NCCS is doubling the power available to its computer room in anticipation of the world's first petascale supercomputer.

Four 13.8 kV transformers are being installed in the facility, and will boost power available to the computer systems from 7.3 to 14 megawatts and the power available for related infrastructure from 2 to 6 megawatts. The added power from the Tennessee Valley Authority, which will be available in December 2007, is necessary because the Cray "Baker" supercomputer, with a peak performance of more than 1,000 trillion calculations a second (1 petaflop), will itself require more than 6 megawatts of power. This is more than two times the power consumed by the center's current premier supercomputer, the Cray XT4 Jaguar system.

### **NCCS Publication Wins Design Award**

*Award is one of many for ORNL, designers*

A publication from the NCCS recently won ORNL an American Graphic Design Award.

The promotional magazine, entitled "15 Years of Leadership Computing," chronicled the history of supercomputing at ORNL from its inception in 1992 to the NCCS's current status as one of the top high-performance computing centers in the world. Over the years

the NCCS has assisted researchers in their efforts to tackle some of science's most pressing questions in areas as diverse as automotive materials and supernova explosions. The award, sponsored by Adobe Systems, Inc., is one of many for ORNL and its graphic designers.